

					GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.					
OM protein - protein search, using sw model					
Run on:	March 1, 2001, 16:09:33 ; Search time 37.5 Seconds				
Perfect score:	382				
Sequence:	1 HDEDEDRGGHSLOOCVORC.....EQEEEOGRGRGWHGGEREE	66	(without alignments)		
Scoring table:	BLOSUM2				
Gapop 10.0 , Gapext 0.5					
Searched:	268485 seqs, 34193795 residues				
Total number of hits satisfying chosen parameters:	268485				
Minimum DB seq length: 0					
Maximum DB seq length: 2000000000					
Post-processing: Minimum Match 0%					
Maximum Match 100%					
Database :					
A_Geneseq_365,*					
1: /SIDS1/gcdata/geneseq/geneseq/geneseq/AA1980.DAT,*					
2: /SIDS1/gcdata/geneseq/geneseq/geneseq/AA1981.DAT,*					
3: /SIDS1/gcdata/geneseq/geneseq/geneseq/AA1982.DAT,*					
4: /SIDS1/gcdata/geneseq/geneseq/geneseq/AA1983.DAT,*					
5: /SIDS1/gcdata/geneseq/geneseq/AA1984.DAT,*					
6: /SIDS1/gcdata/geneseq/geneseq/AA1985.DAT,*					
7: /SIDS1/gcdata/geneseq/geneseq/AA1987.DAT,*					
8: /SIDS1/gcdata/geneseq/geneseq/AA1988.DAT,*					
9: /SIDS1/gcdata/geneseq/geneseq/AA1989.DAT,*					
10: /SIDS1/gcdata/geneseq/geneseq/AA1990.DAT,*					
11: /SIDS1/gcdata/geneseq/geneseq/AA1991.DAT,*					
12: /SIDS1/gcdata/geneseq/geneseq/AA1992.DAT,*					
13: /SIDS1/gcdata/geneseq/geneseq/AA1993.DAT,*					
14: /SIDS1/gcdata/geneseq/geneseq/AA1994.DAT,*					
15: /SIDS1/gcdata/geneseq/geneseq/AA1995.DAT,*					
16: /SIDS1/gcdata/geneseq/geneseq/AA1996.DAT,*					
17: /SIDS1/gcdata/geneseq/geneseq/AA1997.DAT,*					
18: /SIDS1/gcdata/geneseq/geneseq/AA1998.DAT,*					
19: /SIDS1/gcdata/geneseq/geneseq/AA1999.DAT,*					
20: /SIDS1/gcdata/geneseq/geneseq/AA2000.DAT,*					
21: /SIDS1/gcdata/geneseq/geneseq/AA2000.DAT,*					
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.					
Result No.	Score	Query Match Length	DB ID	Description	
1	382	100.0	637 19	Hordeum vulgare an-	
2	138	36.1	593 19	zea mays antimicro-	
3	115	30.1	666 19	Macadamia integrif-	
4	105	27.5	666 19	Macadamia integrif-	
5	93	24.3	590 19	Gossypium hirsutum	
6	91	23.8	625 19	Macadamia integrif-	
7	75	19.6	699 21	Human adipose tiss-	
8	74.5	19.5	1856 20	B. subtilis rib op-	
9	74.5	19.5	1856 21	Polypeptide encode	
10	73	19.1	605 19	Glycine max antimi-	
11	73	19.1	605 20	Soybean bete congri	
12	72	18.8	w62831	Theobroma cacao an-	
				Sequence encoded b-	
				Human ovarian tumo-	
				Drosophila PER B	
				Drosophila PER C	
				Drosophila PER A p	
				HRV8 ORF 21	
				Human metastasis-a-	
				G. max truncated S	
				G. max SBP2 protei-	
				A human trichohyal	
				C. acidovorans gam-	
				Arabidopsis thalia	
				Arabidopsis patho-	
				T. gondii immunoge-	
				Coffee storage pro-	
				Euploites telomeras	
				A human monocyte-m-	
				Yeast transactivatio-	
				Bovine butyrophilic	
				Human prostate tum-	
				Mouse SRY-related	
				Chicken leucocytoz-	
				S. avermitillilis HPP	
				Mouse multiple reg-	
				Human myelin basic	
				Foetal myelin basi-	
				ALIGNMENTS	
				RESULT 1	
				ID W62837 standard; Protein: 637 AA.	
				XX W62837;	
				AC XX	
				DT XX	
				DE XX	
				KW XX	
				OS XX	
				PN XX	
				PD XX	
				XX 02-JUL-1998.	
				XX PF 22-DEC-1997; 97MO-AU00874.	
				PR 20-DEC-1996; 96AU-0004275.	
				XX PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.	
				PA PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;	
				XX WPI; 1998-377279/32.	
				XX PT Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals.	
				XX PS Claim 1; Page 60-62; 96pp; English.	
				CC CC The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian	

CC	animals.		ID	W62828 standard; Protein; 666 AA.
XX			XX	W62828;
SQ	Sequence 637 AA:		AC	W62828;
Query	Match Best Local Similarity 100.0%; Pred. No. 8. 9e-39; Mismatches 0; Indels 0; Gaps 0;	Score 382; DB 19; Length 637; Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	DT	27-OCT-1998 (first entry)
OY	1 HDDEDDRGHSLQCYVORCROPRYSHARCVQECRDOOGHGRHEOEEOGRGRGWH 60	XX DE Macadamia integrifolia antimicrobial protein.	KW	antimicrobial protein; infestation; control.
Db	29 hdeddrgrghsqqcgrcrqerprysharcvqecrddqgaghneqeeggrgwg 88	XX OS Macadamia integrifolia.		
OY	61 EGGEREE 66	XX FH Key	Location/Qualifiers	
Db	89 egeree 94	FT Peptide 1..28 /note= "signal peptide"		
RESULT	2	FT Protein 29..566 /note= "mature protein"		
W62835		FT		
ID	W62835 standard; Protein; 593 AA.	XX		
XX		PN		
AC	W62835;	XX		
XX		PD		
DT	27-OCT-1998 (first entry)	XX		
DE	Zea mays antimicrobial protein.	PF		
XX		XX		
KW	antimicrobial protein; infestation; control.	PR		
XX		XX		
OS	Zea mays.	20-DEC-1996; 96AU-AU00874.		
XX		XX		
PN	W09827805-A1.	PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.		
XX		XX		
PD	02-JUL-1998.	PT Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;		
XX		XX		
PF	22-DEC-1997; 97WO-AU00874.	DR WPI; 1998-377279/32.		
XX		XX		
PR	20-DEC-1996; 96AU-0004275.	PT Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals		
XX		XX		
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.	PS Claim 1; Page 34-36; 96pp; English.		
XX		XX		
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.	CC The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian animals.		
XX		XX		
PT	Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;	SQ Sequence 666 AA;		
XX		XX		
DR	WPI; 1998-377279/32.	Query Match Best Local Similarity 36.1%; Score 138; DB 19; Length 593; Matches 26; Conservative 14; Mismatches 21; Indels 8; Gaps 3;		
XX		Score 115; DB 19; Length 666; Matches 25; Conservative 12; Mismatches 26; Indels 4; Gaps 3;		
PT	Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals	QY 3 DDEDDRGGHSLQCYVORCROPRYSHARCVQECRDOOGHGRHEOEEOGRGRGWH 59		
XX		DB 182 eednkrdpqreyedrrceeqprqhq-qcqrlcreqqhqrggammnpqrgsgry 240		
PS	Claim 1; Page 58-60; 96pp; English.	OY 60 EGGEREE 66		
XX		DB 241 eegeeeq 247		
CC	The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian animals.	SQ Sequence 593 AA;		
CC		RESULT 4		
CC		W62829 ID W62829 standard; Protein; 666 AA.		
XX		XX AC W62829;		
Db	26 dnhhhggksgqcvrc-edrphqrprcleqqreerekrqrsheadrsgg - 82	XX DT 27-OCT-1998 (first entry)		
OY	58 WHGEEREE 66	XX DE Macadamia integrifolia antimicrobial protein.		
Db	83 -ssedreq 90	XX KW antimicrobial protein; infestation; control.		
RESULT	3	OS Macadamia integrifolia.		
W62828		XX FH Key		
		Location/Qualifiers		

FT	Peptide	1...28	/note= "signal peptide"	PT Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals
FT	Protein	XX	/note= "mature protein"	XX
PN	W09827805-A1.	XX		PS Claim 1; Page 49-51; 96pp; English.
PD	02-JUL-1998.	XX		CC The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian animals.
PF	22-DEC-1997;	97WO-AU000874.	XX	CC
PR	20-DEC-1996;	96AU-0004275.	XX	SQ Sequence 590 AA;
XX	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.			
PT	Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;			
XX	WPI; 1998-377279/32.			
DR	N-PADB; v42311.			
XX	Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals			
PT	useful for controlling microbial infestations of plants or mammals			
XX	useful for controlling microbial infestations of plants or mammals			
PS	Claim 1; Page 43-45; 96pp; English.			
XX	The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian animals.			
CC	CC			
CC	The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian animals.			
SQ	Sequence 625 AA;			
Query Match	24.3%	Score 93;	DB 19;	Length 590;
Best Local Similarity	29.4%	Pred. No. 0.0016;		
Matches	20;	Conservative	16;	Mismatches 24;
QY	3 DEDDRRGHSILOOCVQRCKOERPR-YSHARCVQBCRDDQQH--GRHDEEEQGRGKRW 58	Indels 8;	Gaps	
Db	114 eeqeqsqsrqfqeqcqhqchqeqapekkqqcviecrekyqenpwrgereeeaeete-- 171			
QY	59 HGFGEFEE 66			
Db	172 --egeeq 177			
RESULT	6			
WT2830				
ID	W62830 standard; Protein; 625 AA.			
XX				
AC	W62830;			
XX				
DT	27-OCT-1998 (first entry)			
XX				
DE	Macadamia integrifolia antimicrobial protein.			
XX				
KW	antimicrobial protein; infestation; control;			
XX				
OS	Macadamia integrifolia.			
XX				
FH	Key	Location/Qualifiers		
FT	Peptide	1...28		
FT	Protein	/note= "signal peptide"		
FT		29..666		
FT		/note= "mature protein"		
XX				
PN	W09827805-A1.			
XX				
PD	02-JUL-1998.			
XX				
PF	22-DEC-1997;	97WO-AU000874.		
PR	20-DEC-1996;	96AU-0004275.		
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.			
XX				
PT	Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;			
XX	WPI; 1998-377279/32.			
DR	N-PADB; v42316.			
XX	Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals			
PS	Claim 1; Page 43-45; 96pp; English.			
XX	The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian animals.			
CC	CC			
CC	The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian animals.			
SQ	Sequence 625 AA;			
Query Match	23.8%	Score 91;	DB 19;	Length 525;
DR	WPI; 1998-377279/32.			

Best Local Similarity 33.3%; Pred. No. 0.0031; Mismatches 21; Indels 2; Gaps 2;

Matches 18; Conservative 13; Mismatches 21; Indels 2; Gaps 2;

KW Riboflavin; open reading frame; ORF; structural gene; promoter;

XX Vitamin B2; *Bacillus subtilis*; rib operon.

QY 14 DDCVQRQRCRQERYSRSHARCVQECRDDQQHGRHEGEEEOGRGRGWIGEGREE 66  
 XX OS Bacillus subtilis.

Db 154 edcrhqeinqeqprly-qcqrqceqeqrgqhgrrgqglmnpqrgsgryeegeekq 206  
 XX FH Key FT MISC-difference 1..1856  
 FT /note= "residues Xaa are encoded by stop codons"

RESULT 7  
 Y67598 standard; Protein; 699 AA.  
 XX PN US5925538-A.  
 AC Y67598;  
 XX PD 20-JUL-1999.  
 DT 13-JUN-2000 (first entry)  
 XX PR 24-AUG-1998; 98US-0138775.  
 DE Human adipose tissue protein #1.  
 XX PR 11-SEP-1990; 90US-0581048.  
 PR 22-JUN-1989; 89US-0370378.  
 KW Adipose tissue; obesity; diabetes; hyperlipidemia; hypertension; human;  
 PR 21-APR-1992; 92US-0873572.  
 arteriosclerosis; hyperuricemia; sleep apnea syndrome.  
 XX PR 06-FEB-1995; 95US-0384626.  
 OS Homo sapiens.  
 XX PR 24-AUG-1998; 98US-0138775.  
 PA (HOFF ) ROCHE VITAMINS INC.  
 PN XX PA  
 XX PT Erdenerberger T, Hatch RT, Hermann T, Perkins JB;  
 PD Pero JG, Sioma A;  
 XX PI  
 PF 23-JUL-1998; 98JP-0225228.  
 XX DR WPI; 1999-418271/35.  
 PR 23-JUL-1998; 98JP-0225228.  
 XX DR N-PSDB; X81947.  
 XX PT XX  
 PT Use of bacterial strains for the over production of riboflavin  
 XX PS Example 2; Fig 3A-S; 61pp; English.  
 XX CC The invention relates to a method of producing riboflavin that comprises,  
 PT providing a recombinant bacterium containing an exogenously introduced  
 XX bacterial ORF (open reading frame) 3 (or ORFs) structural gene and an  
 CC exogenous promoter; culturing the bacterium, and recovery of the product.  
 PS The method is useful for the production of large quantities of riboflavin  
 XX (vitamin B2). Sequences Y21801-803 are protein sequences translated from  
 CC three different reading frames of B. subtilis rib operon. Y21804-806 are  
 CC protein sequences translated from three different reading frames of the  
 CC complementary sequence.  
 XX SQ Sequence 1856 AA;

Query Match 19.5%; score 74.5; DB 20; Length 1856;  
 Best Local Similarity 22.8%; Pred. No. 1.1; Mismatches 34; Indels 27; Gaps 3;  
 Matches 21; Conservative 10; Mismatches 21; Indels 27; Gaps 3;

QY 2 DDEDDR-----RGGHSLQQCVQRQCRQERYSRSHARCVQECRDDQQHGRHEOE 49  
 XX Db 477 dqhderswhxqkgphgrpgqgeaxkisalyedpsvrnaksqcapxrqdsyqngxqmd 536  
 QY 50 EEQG-----RCRGWIGEGERBE 66  
 Db 537 hvrcktgcsalqentpkhfssrwsrsrqe 568

RESULT 9  
 Y83270 ID 1856 standard; Protein; 1856 AA.  
 XX AC Y83270;  
 XX DT 16-AUG-2000 (first entry)  
 XX DE Polypeptide encoded by rib operon of *Bacillus subtilis*.  
 XX Riboflavin; rib operon; yeast; bacteria; transformation;  
 KW transcription element; gene expression; hair loss; skin inflammation;  
 growth disorder; vision; vision; sight.

RESULT 8  
 Y21802 ID Y21802 standard; Protein; 1856 AA.  
 XX AC Y21802;  
 XX DT 10-SEP-1999 (first entry)  
 XX DE *B. subtilis* rib operon protein translated from reading frame 2.

XX  
OS Bacillus subtilis.

XX  
FH  
FT Key Location/Qualifiers

FT Misc-difference 67 /note= "Encoded by TAA stop codon"

FT Misc-difference 91 /note= "Encoded by TAA stop codon"

FT Misc-difference 92 /note= "Encoded by TAA stop codon"

FT Misc-difference 111 /note= "Encoded by TGA stop codon"

FT Misc-difference 120 /note= "Encoded by TAG stop codon"

FT Misc-difference 133 /note= "Encoded by TGA stop codon"

FT Misc-difference 176 /note= "Encoded by TAA stop codon"

FT Misc-difference 146 /note= "Encoded by TAA stop codon"

FT Misc-difference 151 /note= "Encoded by TAA stop codon"

FT Misc-difference 201 /note= "Encoded by TAA stop codon"

FT Misc-difference 208 /note= "Encoded by TAA stop codon"

FT Misc-difference 234 /note= "Encoded by TAA stop codon"

FT Misc-difference 245 /note= "Encoded by TGA stop codon"

FT Misc-difference 253 /note= "Encoded by TAA stop codon"

FT Misc-difference 269 /note= "Encoded by TAA stop codon"

FT Misc-difference 295 /note= "Encoded by TAG stop codon"

FT Misc-difference 298 /note= "Encoded by TAA stop codon"

FT Misc-difference 323 /note= "Encoded by TAA stop codon"

FT Misc-difference 341 /note= "Encoded by TAA stop codon"

FT Misc-difference 344 /note= "Encoded by TAA stop codon"

FT Misc-difference 356 /note= "Encoded by TAA stop codon"

FT Misc-difference 358 /note= "Encoded by TAA stop codon"

FT Misc-difference 413 /note= "Encoded by TAA stop codon"

FT Misc-difference 469 /note= "Encoded by TGA stop codon"

FT Misc-difference 486 /note= "Encoded by TAA stop codon"

FT Misc-difference 501 /note= "Encoded by TGA stop codon"

FT Misc-difference 523 /note= "Encoded by TGA stop codon"

FT Misc-difference 532 /note= "Encoded by TAA stop codon"

FT Misc-difference 586 /note= "Encoded by TGA stop codon"

FT Misc-difference 593 /note= "Encoded by TGA stop codon"

FT Misc-difference 596 /note= "Encoded by TAA stop codon"

FT Misc-difference 632 /note= "Encoded by TAA stop codon"

FT Misc-difference 640 /note= "Encoded by TGA stop codon"

FT Misc-difference 681 /note= "Encoded by TGA stop codon"

FT Misc-difference 695 /note= "Encoded by TAA stop codon"

FT Misc-difference 710 /note= "Encoded by TGA stop codon"

FT Misc-difference 717 /note= "Encoded by TGA stop codon"

FT Misc-difference 731 /note= "Encoded by TGA stop codon"

FT Misc-difference 747 /note= "Encoded by TGA stop codon"

FT Misc-difference 755 /note= "Encoded by TAA stop codon"

FT Misc-difference 764 /note= "Encoded by TAA stop codon"

FT Misc-difference 775 /note= "Encoded by TAG stop codon"

FT Misc-difference 780 /note= "Encoded by TGA stop codon"

FT Misc-difference 804 /note= "Encoded by TGA stop codon"

FT Misc-difference 807 /note= "Encoded by TAA stop codon"

FT Misc-difference 813 /note= "Encoded by TAA stop codon"

FT Misc-difference 860 /note= "Encoded by TAA stop codon"

FT Misc-difference 862 /note= "Encoded by TAA stop codon"

FT Misc-difference 875 /note= "Encoded by TAA stop codon"

FT Misc-difference 877 /note= "Encoded by TAA stop codon"

FT Misc-difference 882 /note= "Encoded by TAA stop codon"

FT Misc-difference 910 /note= "Encoded by TAA stop codon"

FT Misc-difference 916 /note= "Encoded by TGA stop codon"

FT Misc-difference 942 /note= "Encoded by TAA stop codon"

FT Misc-difference 975 /note= "Encoded by TAA stop codon"

FT Misc-difference 976 /note= "Encoded by TGA stop codon"

FT Misc-difference 981 /note= "Encoded by TGA stop codon"

FT Misc-difference 997 /note= "Encoded by TGA stop codon"

FT Misc-difference 1013 /note= "Encoded by TAA stop codon"

FT Misc-difference 1020 /note= "Encoded by TGA stop codon"

FT Misc-difference 1026 /note= "Encoded by TGA stop codon"

FT Misc-difference 1028 /note= "Encoded by TGA stop codon"

FT Misc-difference 1044 /note= "Encoded by TGA stop codon"

FT Misc-difference 1070 /note= "Encoded by TGA stop codon"

FT Misc-difference 1099 /note= "Encoded by TGA stop codon"

FT Misc-difference 1102  
 FT /note= "Encoded by TGA stop codon"  
 FT Misc-difference 1114  
 FT /note= "Encoded by TGA stop codon"  
 FT Misc-difference 1118  
 FT /note= "Encoded by TGA stop codon"  
 FT Misc-difference 1127  
 FT /note= "Encoded by TGA stop codon"  
 FT Misc-difference 1130  
 FT /note= "Encoded by TGA stop codon"  
 FT Misc-difference 1143  
 FT /note= "Encoded by TAA stop codon"  
 FT Misc-difference 1159  
 FT /note= "Encoded by TGA stop codon"  
 FT Misc-difference 1161  
 FT /note= "Encoded by TGA stop codon"  
 FT Misc-difference 1167  
 FT /note= "Encoded by TGA stop codon"  
 FT Misc-difference 1172  
 FT /note= "Encoded by TAA stop codon"  
 FT Misc-difference 1179  
 FT /note= "Encoded by TGA stop codon"  
 FT Misc-difference 1212  
 FT /note= "Encoded by TGA stop codon"  
 FT Misc-difference 1230  
 FT /note= "Encoded by TGA stop codon"  
 FT Misc-difference 1255  
 FT /note= "Encoded by TAA stop codon"  
 FT Misc-difference 1260  
 FT /note= "Encoded by TAA stop codon"  
 FT Misc-difference 1267  
 FT /note= "Encoded by TGA stop codon"  
 FT Misc-difference 1273  
 FT /note= "Encoded by TGA stop codon"  
 FT Misc-difference 1305  
 FT /note= "Encoded by TAA stop codon"  
 FT Misc-difference 1313  
 FT /note= "Encoded by TGA stop codon"  
 FT Misc-difference 1334  
 FT /note= "Encoded by TAA stop codon"  
 FT Misc-difference 1517  
 FT /note= "Encoded by TAA stop codon"  
 FT Misc-difference 1529

RESULT 11  
 Y40999  
 ID Y40999 standard; protein; 605 AA.  
 XX  
 AC Y40999;  
 XX  
 DT 06-DEC-1999 (first entry)  
 XX  
 DE Soybean beta-conglycinin protein sequence.  
 XX  
 KW Peanut; allergen; Ara H 1; IGE; immunoglobulin E; epitope; Ara h 3;  
 KW allergic reaction; soybean; beta-conglycinin.  
 OS Glycine max.  
 PN W09943961-A1.  
 XX  
 PD 16-SEP-1999.  
 XX  
 PF 12-MAR-1999; 99WO-US05494.  
 PR 12-MAR-1998; 98US-0077763.  
 PR 11-MAR-1999; 99US-007773.  
 XX  
 PA (UYAR-) UNIV ARKANSAS.  
 XX  
 PI Burks W, Helm RM, Cockrell G, Baun GA, Stanley JS, Shin DS;  
 XX  
 DR Burks W, Helm RM, Cockrell G, Baun GA, Stanley JS, Shin DS;  
 XX  
 PI Sampson H, Compadre CM, Huang SK, Maleki SJ, Kopper RA;  
 XX  
 DR Burks W, Helm RM, Cockrell G, Baun GA, Stanley JS, Shin DS;  
 XX  
 PR Tertiary structure of peanut allergen Ara h 1 for protection of a host  
 XX  
 PS Disclosure; Fig 33A-B; 195pp; English.

Query Match 19.5%; Score 74.5; DB 21; Length 1856;  
 Best Local Similarity 22.8%; PRed. No. 1,1; Mismatches 21; Gaps 21; Conservativeness 10; Indels 3; Matches 20; Mismatches 6; Indels 26; Gaps 3;

QY 15 QCVQRQERRPRY---SHARC-----VQG-----RDPQQQHGRHEQ 48  
 Db 37 kclqscnserdsyrsnqacharcnlikvekeecegeiprprprpqhperepgqpekeed 96  
 QY 49 EEEBQGR 54  
 Db 97 edeqpr 102

Query Match 19.1%; Score 73; DB 19; Length 605;  
 Best Local Similarity 30.3%; Pred. No. 0,47; Mismatches 14; Indels 26; Gaps 3; Matches 20; Conservativeness 10; Indels 26; Gaps 3;

QY 15 QCVQRQERRPRY---SHARC-----VQG-----RDPQQQHGRHEQ 48  
 Db 37 kclqscnserdsyrsnqacharcnlikvekeecegeiprprprpqhperepgqpekeed 96

RESULT 11  
 Y40999  
 ID Y40999 standard; protein; 605 AA.  
 XX  
 AC Y40999;  
 XX  
 DT 06-DEC-1999 (first entry)  
 XX  
 DE Soybean beta-conglycinin protein sequence.  
 XX  
 KW Peanut; allergen; Ara H 1; IGE; immunoglobulin E; epitope; Ara h 3;  
 KW allergic reaction; soybean; beta-conglycinin.  
 OS Glycine max.  
 PN W09943961-A1.  
 XX  
 PD 16-SEP-1999.  
 XX  
 PF 12-MAR-1999; 99WO-US05494.  
 PR 12-MAR-1998; 98US-0077763.  
 PR 11-MAR-1999; 99US-007773.  
 XX  
 PA (UYAR-) UNIV ARKANSAS.  
 XX  
 PI Burks W, Helm RM, Cockrell G, Baun GA, Stanley JS, Shin DS;  
 XX  
 DR Burks W, Helm RM, Cockrell G, Baun GA, Stanley JS, Shin DS;  
 XX  
 PR Tertiary structure of peanut allergen Ara h 1 for protection of a host  
 XX  
 PS Disclosure; Fig 33A-B; 195pp; English.

RESULT 10  
 W62838  
 ID W62838 standard; Protein; 605 AA.  
 XX  
 AC W62838;  
 XX  
 DT 27-OCT-1998 (first entry)  
 XX  
 DE Glycine max antimicrobial protein.  
 XX  
 KW antimicrobial protein; infestation; control.  
 XX  
 OS Glycine max.  
 XX  
 PN W09827805-A1.



DT 10-APR-2000 (first entry)  
 XX Human ovarian tumor EST fragment encoded protein 79.  
 DE KW  
 XX Expressed sequence tag; EST; human; ovarian tumor; anticancer;  
 KW gene therapy; treatment.  
 KW OS Homo sapiens.  
 XX PN DE1981/557-A1.  
 XX PD 21-OCT-1999.  
 XX PF 09-APR-1998; 98DE-1017557.  
 XX PR 09-APR-1998; 98DE-1017557.  
 XX PA (META-) METAGEN GES GENOMFORSCHUNG MBH.  
 XX PT Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;  
 XX WPI; 1999-591920/51.  
 DR N-PSDB; Z77484.  
 XX PT New nucleic acid sequences expressed in ovarian, and some other, cancer tissues, and derived polypeptides, for treatment of ovarian cancer and identification of therapeutic agents -  
 PT Claim 25; Page 276; 310pp; German.  
 XX PS This invention describes novel nucleic acid (cDNA) sequences (A) which have anticancer activity and are highly expressed in ovarian tumor tissue (and som also in testis and breast cancer tissue). The products of the invention can be used for gene therapy. (A) are used (1) for recombinant expression of polypeptides (B) and (ii) to isolate complete genes. (B) are used (i) to identify agents suitable for treatment of ovarian cancer; (ii) directly for treating this form of cancer (including expression from gene therapy vectors) and (iii) for generation of specific antibodies. (A) are identified by assembling ESTs (expressed sequence tags) from a particular tissue type before comparison of expression patterns. This allows a significantly longer fragment of the gene to be revealed, so should reduce the number of failures associated with the fact that ESTs from different libraries may represent different parts of the same unknown gene, distorting the estimated frequency of occurrence in a particular tissue. Y76505-Y76638 represent protein fragments, encoded by the human ovarian tumor cDNA library derived EST CC fragments, represented in Z77450-Z77572.

XX SQ Sequence 156 AA:

Query Match 18.3%; Score 70; DB 20; Length 156;  
 Best Local Similarity 28.6%; Pred. No. 0.24; Mismatches 18; Conservative 11; Mismatches 28; Indels 6; Gaps 2;

QY 7 RRGGHSLQQCVORCRODRPRYSH-----RCVQECRDRDQQHGRHRHEEEEOGRGRGWHGEG 62  
 Db 70 rsgskartpqifrlqqqlqrfrghgcevprcwliqarehpgggqeaeegq--eeg 127

Oy 63 ERE 65  
 Db 128 qee 130

RESULT 15  
 Y32218  
 ID Y32218 standard; Protein; 1122 AA.  
 XX AC Y32218;  
 XX DT 15-FEB-2000 (first entry)  
 XX DE Drosophila PER B protein.

XX PER B; period; perl gene; transcription factor; circadian rhythm;  
 KW KW jet lag; sleep disorder; depression; seasonal affective disorder;  
 KW fertility; therapy.  
 XX OS Drosophila melanogaster.  
 XX PN WO9957137-A1.  
 XX PD 11-NOV-1999.  
 XX PF 06-MAY-1999; 99WO-US10072.  
 XX PR 07-MAY-1998; 98US-0084610.  
 XX PA (HARD ) HARVARD COLLEGE.  
 XX PI Weitz CJ, Gekakis N, Staknis D;  
 DR XX WPI; 2000-052938/04.  
 XX PT Novel heterodimeric composition for identifying modulators used in diagnosing and treating circadian clock disruption disorders -  
 XX PS Disclosure; Fig 31; 96pp; English.  
 CC This sequence represents the Drosophila melanogaster PER protein PER B. Splice variants are PER A (see Y32217) and PER C (see Y32219). PER forms a heteromeric complex together with TIM (see Y32221) protein and has a biological activity which inhibits transcription of the per1 gene when the CLOCK protein is present in combination with BMAL1 protein (see Y32209). The invention is based on the discovery of the transcriptional mechanism regulating genes responsible for the establishment and/or maintenance of the circadian clock, and provides an assay for novel drugs aimed at restoration of a normal circadian cycle, the drugs being modulators of BMAL1-CLOCK-mediated transcription of E-box-linked genes. The drugs are used to treat conditions such as jet lag, sleep disorders, depression (seasonal affective disorder) and infertility. The invention also provides BMAL1 and CLOCK proteins with which to stimulate the transcription of an E-box-linked gene which regulates the circadian clock.  
 CC XX SQ .Sequence 1122 AA;

Query Match 18.3%; Score 70; DB 21; Length 1122;  
 Best Local Similarity 36.2%; Pred. No. 2.2; Mismatches 21; Conservative 9; Mismatches 14; Indels 14; Gaps 4;

QY 22 QERRPRYSHARCVQE-CRDQQ-QQHGRH-----EDEEEOGRRGRGWHGEGREE 66  
 Db 103 eekprpsgtgceeqqicreldqgqgedhsqeqaleqlgggeeqeds-gseseadrve 159

Search completed: March 1, 2001, 16:09:35  
 Job time: 11327 sec